

## Singapore Management University Institutional Knowledge at Singapore Management University

Research Collection School Of Economics

School of Economics

9-2014

# Intraday periodicity adjustments of transaction duration and their effects on high-frequency volatility estimation

Yiu Kuen TSE

Singapore Management University, [yktse@smu.edu.sg](mailto:yktse@smu.edu.sg)

Yingjie DONG

Singapore Management University

**DOI:** <https://doi.org/10.1016/j.jempfin.2014.04.004>

Follow this and additional works at: [https://ink.library.smu.edu.sg/soe\\_research](https://ink.library.smu.edu.sg/soe_research)



Part of the [Econometrics Commons](#), and the [Finance Commons](#)

### Citation

TSE, Yiu Kuen and DONG, Yingjie. Intraday periodicity adjustments of transaction duration and their effects on high-frequency volatility estimation. (2014). *Journal of Empirical Finance*. 28, (4), 352-361. Research Collection School Of Economics.

**Available at:** [https://ink.library.smu.edu.sg/soe\\_research/1572](https://ink.library.smu.edu.sg/soe_research/1572)

This Journal Article is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [libIR@smu.edu.sg](mailto:libIR@smu.edu.sg).

# Intraday Periodicity Adjustments of Transaction Duration and Their Effects on High-Frequency Volatility Estimation

Yiu-Kuen Tse

School of Economics, Singapore Management University

Yingjie Dong

School of Economics, Singapore Management University

June 2012

**Abstract:** Intraday periodicity is an important feature of high-frequency financial data. In this paper we study two methods of adjusting for intraday periodicity: the Duration Adjustment (DA) method and the Time Transformation (TT) method, and examine their effects on the estimation of intraday volatility. Using the ACD-ICV method of Tse and Yang (2012) to estimate high-frequency volatility, we find that daily volatility estimates are not sensitive to intraday periodicity adjustments. However, intraday volatility is found to have a stronger U-shaped volatility smile when intraday periodicity adjustment is applied to the data than when it is not. The TT method has advantages over the DA method due to its stronger theoretical underpinning and easier implementation in practice.

**JEL Codes:** C410, G120

**Keywords:** Autoregressive conditional duration model, Intraday volatility, Time transformation, Transaction data

**Acknowledgment:** The authors gratefully acknowledge research support from the Center of Financial Econometrics at the Sim Kee Boon Institute for Financial Economics, Singapore Management University.

**Corresponding Author:** Yiu-Kuen Tse, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, email: [yktse@smu.edu.sg](mailto:yktse@smu.edu.sg).

# 1 Introduction

A well known problem in analyzing high-frequency financial data is the stylized fact of intraday periodicity: Trading activities are usually higher at the beginning and close of the trading day than around lunch time. This trading pattern induces the average transaction duration to exhibit an inverted U-shape over the trading day. As argued by Andersen and Bollerslev (1997) in their study on foreign exchange and equity markets, intraday periodicity has a strong impact on the autocorrelation pattern of the absolute intraday returns. They point out that volatility over different intervals of the same calendar-time length at different times of the day may differ due to their differences in trading activities.

Recognizing that duration is not exogenous to the price process, Engle and Russell (1998) introduce the Autoregressive Conditional Duration (ACD) model. They propose to correct for the intraday duration pattern prior to fitting the ACD model to the data. Specifically, they apply the Duration Adjustment (DA) method to adjust for transaction duration as

$$\tilde{x}_{i+1} = \frac{x_{i+1}}{\phi(t_i)}, \quad (1)$$

where  $t_i$  is the calendar time of occurrence of the  $i$ th trade,  $x_{i+1} = t_{i+1} - t_i$  is the duration of the  $(i + 1)$ th trade in calendar time,  $\tilde{x}_{i+1}$  is the diurnally adjusted duration of the  $(i + 1)$ th trade and  $\phi(\cdot)$  is the diurnal adjustment factor with its argument usually taken as the calendar time  $t_i$ .

There are several ways to estimate the diurnal factor  $\phi(\cdot)$  in equation 1. Bauwens and Giot (2000) define the diurnal factor  $\phi(\cdot)$  as the expected duration conditional on the time of the day for each day of the week. The expected duration is computed by averaging the durations over 30-minute intervals. Cubic splines are then used to smooth the time-of-the-day function. Other methods include those proposed by Rodríguez-Poo *et al.* (2007), Dufour and Engle (2000), Tsay (2002) and Drost and Werker (2004). The implementation of the DA method, however, is dependent on the specific smoothing method applied to obtain the diurnal factor. Also, for the purpose of computing adjusted duration it is not sure which time point within the interval should be used for adjustment. This criticism is especially relevant for illiquid stocks for which the transaction duration may be quite long.

In this paper we consider an alternative strategy to correct for intraday periodicity, call the Time Transformation (TT) method. The theoretical underpinning of the TT method is that if there were no intraday differences in trading activities, we would expect the transactions to be evenly spread out throughout the trading day. Under the TT method a time transformation function is determined using empirical data so that the transactions are evenly observed throughout the day under the transformed time. We distinguish between calendar time and diurnally transformed time. Under the diurnally transformed time, the intraday periodicity which induces the differences in the average intraday transaction duration should disappear.

To assess the empirical implications of these adjustment methods, we study their effects on intraday volatility estimation. Following the method proposed by Tse and Yang (2012) for the estimation of

high-frequency volatility, called the ACD-ICV method, we examine the effects of the use of the DA and TT methods to adjust for intraday transaction durations on the estimation of intraday volatility. Our findings are as follows. First, for the estimation of daily volatility, whether or not the durations are diurnally adjusted makes little difference. Second, to estimate intraday volatility over half-hour intervals, correcting for intraday periodicity using either the DA or TT method produces more prominent U-shaped volatility smiles than not adjusting for intraday periodicity. Third, the periodicity observed in the serial correlation of absolute returns over 5-minute intervals is much reduced when return is standardized by daily volatility scaled by the diurnally transformed duration or 5-minute intraday volatility estimate. Finally, our simulation results show that the ACD model with adjustment using the TT method is able to replicate the stylized fact of an inverted U-shaped average-duration curve.

The balance of the paper is as follows. Section 2 describes the data and their main features. In Section 3 we outline the DA and TT methods, and discuss their implementation in practice. Section 4 reports some empirical results of the daily and intraday volatility estimates with and without intraday periodicity adjustment.

## 2 High-Frequency Financial Data and Intraday Periodicity

The transaction data used in this paper were extracted and compiled from the New York Stock Exchange (NYSE) Trade and Quote (TAQ) Database provided through the Wharton Research Data Services. We selected ten stocks from each of the top one-third and bottom one-third of the 500 component stocks of the S&P500 index ranked by market capitalization, and call these large-cap and small-cap stocks, respectively.<sup>1</sup> Data from the Consolidated Trade (CT) file, including the date, trading time, price and number of shares traded were extracted for each stock, over the period January 3, 2005 through December 31, 2007. We dropped some abnormal trading days that may contaminate the results. These include trading days with first trade after 11:00 or last trade before 13:30. We also deleted days with no price events in the last two hours of the day, which may be due to some special incidents. The data were then filtered before the empirical analysis. We deleted entries with corrected trades, i.e., trades with a correction indicator. Only data with a regular sale condition and code E and F were selected.<sup>2</sup>

Table 1 presents some summary statistics of the selected stocks. Trades with the same time stamp can be treated as separate trades or one trade, and we report summary statistics for both assumptions.<sup>3</sup> It can be seen that all stocks in our sample are very actively traded. In particular, the large-cap stocks are more frequently traded than the small-cap stocks. For the large-cap stocks the average number of transactions per day ranges from 4746 to 8264 and the average duration per trade ranges from 2.83 seconds to 4.93 seconds, when multiple trades are treated as one trade. On the other hand, when

---

<sup>1</sup>This terminology is adopted for convenience, although we should note that all stocks in this study are in the S&P500 index.

<sup>2</sup>These screening procedures were adopted from Barndorff-Nielsen *et al.* (2009).

<sup>3</sup>When there are multiple trades at the same stamped time, we calculate the price at the stamped time as the average of the prices weighted by trade volume.

multiple trades are treated as separate trades, the average number of transactions per day increases substantially and ranges from 7360 to 14591, while the average duration per trade ranges from 1.6 seconds to 3.18 seconds. While the number of trades with the same time stamp are quite abundant for large-cap stocks, they are less frequent for small-cap stocks. We would also mention that there is some regularity in the occurrence of multiple trades. For example, many multiple trades occur on the hour. In particular, there are many multiple trades at time stamps 10:00 and 15:45.

An important feature of high-frequency financial data is the prevalence of intraday periodicity in transaction activities. Typically, trading activities are higher at the beginning of the trading day due to the opening auction. Also, when there is overnight macroeconomic news or company news release, traders may actively engage in transactions at the opening of the trading day in order to benefit from the news. On the other hand, trading activities may be high at the end of the trading day as some traders may wish to close their positions before the end of the trading session. In contrast to these high trading activities, trading around lunch time may be relatively inactive. Thus, the intraday pattern of the average trade duration may exhibit an inverted U-shape over the trading day. Likewise, the average number of trades at each second over the trading day may exhibit a U-shape. This phenomenon has been well documented in the literature (see Engle and Russell (1998) and Giot (2005).)

Figure 1 plots the total number of trades of IBM (a large-cap stock) and BMS (a small-cap stock) at each second from 9:30 to 16:00 over all trading days in the sample, with multiple trades treated as separate trades. It can be seen that there is significant intraday periodicity: The number of trades at each second has a clear U-shape, although for the small-cap stock BMS trading activities pick up gradually after market opens and then declines towards lunch time before ending at a high when market closes. Figure 2 plots the average trade durations from 9:30 to 16:00 for these two stocks, with smoothing performed using cubic spline (see Appendix A.1 for the details). It can be seen that the average trade duration as a function of calendar time exhibits an inverted U-shape, although for the small-cap BMS stock the average duration drops after market opens before peaking at lunch time.<sup>4</sup>

### 3 Diurnal Adjustment for Intraday Periodicity

We now describe two methods of diurnal adjustment to account for intraday periodicity, namely, the Duration Adjustment (DA) method and the Time Transformation (TT) Method.

#### 3.1 Duration Adjustment Method

The DA method involves adjusting the raw duration using a diurnal factor to obtain the diurnally adjusted duration, as given in equation (1). To specify the diurnal factor  $\phi(\cdot)$  many studies use the

---

<sup>4</sup>To save space, only the graphs for IBM and BMS are presented. Results for other stocks are very similar, and they are presented in the accompanying Appendix of Figures. In computing the average duration in Figure 2 we treat multiple trades as one trade. Treating multiple trades as one trade is equivalent to merging observations with zero duration. This practice has been adopted by some researchers. For example, Wu (2011) deletes observations with zero duration before computing the serial correlation of duration.

regression method with linear spline or cubic spline method proposed by Engle and Russell (1998). While there are quite a few variations in the literature in the estimation of  $\phi(\cdot)$ , we follow the method proposed by Bauwens and Giot (2000), for which the focus is on estimating the expected duration conditional on the time of the day, with some slight modifications. Although Bauwens and Giot (2000) allow the intraday diurnal factor  $\phi(\cdot)$  to be different for each day of the week, we impose the equality restriction for simplicity. We calculate the average trade duration over each 30-minute interval of the trading day. As opening and closing periods are usually more active due to the opening auction and closing effect, we compute the average durations over 15-minute intervals at the opening and close of the trading day. The average durations are used as knots at the mid-points of their respective intervals and the average-duration function over the trading day is computed using the cubic spline smoothing method. We then standardize the average-duration function to obtain the diurnal factor  $\phi(\cdot)$ . The details of the procedure are summarized in Appendix A.1.

While the DA method is commonly used in the literature, it has some important drawbacks. First, the average duration in each interval is a local measure. There is a choice between taking longer intervals (so that there are more observations in each interval) versus shorter intervals (so that the knots are more precisely located). Obviously, differences in the length of the intervals and location of the knots may produce different results for the diurnal factor. Second, given a raw duration, it is not clear which time point within the interval should be taken to evaluate the diurnal factor. As the diurnal factor function takes an inverted U-shape, if the start time of the interval is taken for adjustment, which is the usual practice, the adjustment may be understated for trades before lunch time and overstated for trades after lunch time. Hence, a systematic bias may be introduced in the diurnal adjustment.

### 3.2 Time Transformation Method

We now describe the use of the time transformation method. Let  $t$  denote the calendar time and  $\tilde{t}$  denote the corresponding diurnally transformed time. We assume  $Q^*(t)$  to be the theoretical average proportion of trades in a day up to time  $t$ . In the context of the NYSE, as one trading day has 6.5 hours ( $6.5 \times 60 \times 60 = 23400$  seconds), we measure time in second so that  $0 \leq t \leq 23400$ . Under the assumption of no intraday periodicity the unconditional distribution of trades should be evenly distributed throughout the trading day. Hence, we define the diurnal time transformation by

$$\frac{\tilde{t}}{23400} = Q^*(t). \quad (2)$$

As  $Q^*(t)$  is unobservable, we estimate it empirically from the transaction data, the details of which can be found in Appendix A.2. We denote the empirical estimate of  $Q^*(t)$  by  $Q(t)$ , so that given a calendar-time point  $t$ , the corresponding diurnally transformed time is  $\tilde{t} = 23400Q(t)$ . Conversely, given a diurnally transformed time  $\tilde{t}$ , the corresponding calendar time is

$$t = Q^{-1}\left(\frac{\tilde{t}}{23400}\right), \quad (3)$$

where  $Q^{-1}(\cdot)$  is the inverse function of  $Q(\cdot)$ .

Wu (2011) proposes a similar method independently for diurnally adjusting for intraday periodicity through time transformation. He shows that under an appropriately chosen time transformation function ( $Q^*(t)$  in our context), the resulting process has a first-order stationary conditional intensity. He derives some asymptotic properties for the empirical estimates of the transformation function and demonstrates in a simulation study that the diurnally transformed duration is free from serial correlation. In this paper, we treat the TT method as an empirical procedure to correct for intraday periodicity and focus on the effects of this adjustment on the estimation of intraday volatility.<sup>5</sup>

Given any two calendar-time points  $t_i < t_j$ , the diurnally adjusted duration between these two time points is  $\tilde{t}_j - \tilde{t}_i = 23400 [Q(t_j) - Q(t_i)]$ . Likewise, given any two diurnally transformed time points  $\tilde{t}_i < \tilde{t}_j$ , the corresponding duration in calendar time is

$$Q^{-1}\left(\frac{\tilde{t}_j}{23400}\right) - Q^{-1}\left(\frac{\tilde{t}_i}{23400}\right). \quad (4)$$

Figure 3 presents the plots of the diurnally transformed time  $\tilde{t}$  against the calendar time  $t$  for the IBM and BMS stocks. The solid lines are the plots of  $23400Q(t)$ , computed by treating multiple trades as separate trades. Thus, if there was no intraday periodicity,  $23400Q^*(t)$  should follow the 45 degree line.<sup>6</sup> Note that the  $Q(t)$  function constructed is strictly monotonically increasing, so that  $t$  and  $\tilde{t}$  form a one-to-one correspondence. Under the TT method, we first transform the calendar time of occurrence of trade to diurnally adjusted time. We then calculate the transaction duration using the diurnally adjusted time and fit the ACD model to these duration data. Due to the one-to-one correspondence, diurnally transformed time and duration can also be easily converted to calendar time and duration.

There are some advantages of the TT method over the DA method. First, the  $Q(t)$  function is easy to compute and it depends on all data in the sample. This is in contrast to the DA method, which depends on the choice of the knots. Second, the definition of diurnally adjusted duration is natural. This removes the ambiguity in the choice of the time point within the transaction interval for the application of the diurnal factor. Third, the switch between calendar time and diurnally adjusted time can be performed easily. This facilitates simulation using models in one measure of duration (e.g., ACD model for diurnally adjusted duration) to draw implications for the market in another measure of time (e.g., implications for the market in calendar time). Recently, Dionne *et al.* (2009) suggest using simulation method to estimate the intraday value at risk (IVaR). While the time interval specified may be in calendar time, the duration model estimated may be for diurnally adjusted data. The TT method will be convenient to use as the calendar time and diurnally adjusted time can be easily converted from one another.

---

<sup>5</sup>Wu's (2011) transaction data in his simulation study are generated from *unconditional* Poisson distributions with time varying intensity. In addition, he deletes data from the first 20 minutes and transactions with zero duration (i.e., multiple trades occurring at the same time stamp are treated as one trade). Thus, his results should not be used as evidence against the conditional dependence of transaction duration as captured by the ACD model. Unlike Wu (2011), we do not propose to use the TT method as a correction for duration serial correlation.

<sup>6</sup>The vertical axis in Figure 3 is converted from second to transformed time from 9:30 to 16:00 for ease of reference.

## 4 Diurnal Adjustment and Intraday Volatility Estimation

To examine the empirical implications of the methods of adjusting for intraday periodicity we consider the effects of these adjustments on the estimation of intraday volatility.<sup>7</sup> We use the ACD-ICV method for estimating high-frequency volatility proposed by Tse and Yang (2012), and compare the volatility estimates when the data are (a) not adjusted for intraday periodicity, (b) adjusted for intraday periodicity using the TT method, and (c) adjusted for intraday periodicity using the DA method. In what follows we first briefly summarize the ACD-ICV method, after which we report the empirical results on the NYSE data.

### 4.1 Intraday Volatility Estimation using the ACD-ICV Method

The ACD-ICV method samples observations from transaction data based on a threshold price range  $\delta$ . A price event is said to occur if the logarithmic stock price first moves by an amount  $\delta$  or more, whether upwards or downwards. The waiting time for the price event to occur is called the price duration. Let  $t_0$  be the beginning time of a period, and  $t_1, t_2, \dots, t_N$  be subsequent times of occurrence of price events, so that the price duration of the  $i$ th trade is  $x_i = t_i - t_{i-1}$ . We denote  $\Phi_i$  as the information set upon the transaction at time  $t_i$ , which is assumed to consist of lagged price durations, and define  $\psi_i = E(x_i | \Phi_{i-1})$  as the conditional expectation of the price duration. To model the expected duration we adopt the power ACD (PACD) model (see Fernandes and Grammig (2006)) defined by

$$\psi_i^\lambda = \omega + \alpha x_{i-1}^\lambda + \beta \psi_{i-1}^\lambda, \quad (5)$$

where  $\omega, \alpha, \beta > 0$ . The integrated conditional variance (ICV) is given by

$$\text{ICV} = \delta^2 \sum_{i=0}^{N-1} \frac{t_{i+1} - t_i}{\psi_{i+1}}. \quad (6)$$

Tse and Yang (2012) estimate the parameters in equation (5) with  $\epsilon_i = x_i/\psi_i$  assumed to be i.i.d. standard exponential. Estimates of  $\psi_i$  are then substituted into equation (6), resulting in the ACD-ICV estimate of the variance in the interval  $(t_0, t_N)$ , which may be one trading day or a subinterval of a trading day, such as one hour.

In this paper we consider three alternatives of estimating daily or intraday volatility using the ACD-ICV method. First, we compute the price duration using unadjusted calendar time, and call this the raw duration. We estimate the PACD model using the raw duration, with resulting estimates of  $\psi_i$  denoted by  $\hat{\psi}_i$  and the ICV computed as

$$\text{ICV} = \delta^2 \sum_{i=0}^{N-1} \frac{x_{i+1}}{\hat{\psi}_{i+1}} = \delta^2 \sum_{i=0}^{N-1} \frac{t_{i+1} - t_i}{\hat{\psi}_{i+1}}. \quad (7)$$

---

<sup>7</sup>Research on intraday movements of equity prices has recently attracted much interest due to its implications for market participants such as day traders and market makers, as illustrated by the works of Giot (2005) and Dionne *et al.* (2009).



We call this method M1, which uses raw duration without diurnal adjustment. Next, we take account of intraday periodicity using the TT method and compute the diurnalized duration as

$$\tilde{x}_{i+1} = \tilde{t}_{i+1} - \tilde{t}_i. \quad (8)$$

We estimate the PACD model using the TT diurnalized duration, with resulting estimates of  $\psi_i$  denoted by  $\tilde{\psi}_i$  and the ICV computed as

$$\text{ICV} = \delta^2 \sum_{i=0}^{N-1} \frac{\tilde{x}_{i+1}}{\tilde{\psi}_{i+1}} = \delta^2 \sum_{i=0}^{N-1} \frac{\tilde{t}_{i+1} - \tilde{t}_i}{\tilde{\psi}_{i+1}}. \quad (9)$$

We call this method M2, which uses diurnalized duration based on the TT method. Finally, we denote the diurnally adjusted duration using the DA method by  $\check{x}_i$ , which is given by

$$\check{x}_{i+1} = \frac{t_{i+1} - t_i}{\phi(t_i^M)} = \frac{x_{i+1}}{\phi(t_i^M)}, \quad (10)$$

where  $t_i^M = (t_i + t_{i+1})/2$  is the mid-point of the interval  $(t_i, t_{i+1})$ .<sup>8</sup> We estimate the PACD model using the diurnalized duration computed from the DA method, with resulting estimates of  $\psi_i$  denoted by  $\check{\psi}_i$  and the ICV computed as

$$\text{ICV} = \delta^2 \sum_{i=0}^{N-1} \frac{\check{x}_{i+1}}{\check{\psi}_{i+1}} = \delta^2 \sum_{i=0}^{N-1} \frac{t_{i+1} - t_i}{\check{\psi}_{i+1} \phi(t_i^M)}. \quad (11)$$

We call this method M3, which uses diurnalized duration based on the DA method.

## 4.2 Empirical Results

We compute the daily volatility estimates of the 20 NYSE stocks in our sample using the three methods: M1, M2 and M3. We report in detail the case when M2 is based on the convention that multiple trades are treated as separate trades and M3 is based on the convention that multiple trades are treated as one trade. Table 2 reports the results for the daily volatility estimates, giving the mean of each estimate, and the mean absolute deviation (MAD) and the root mean-squared deviation (RMSD) between pairs of daily estimates. The daily volatility estimates are expressed in annualized return standard deviation in percentage. The results show that the daily volatility estimates of the three methods are very similar. In terms of the MAD, the differences between M2 and M3 (i.e., two methods with adjustment for intraday periodicity) are very small (less than 0.1% for all stocks except for CVX among the large-cap stocks and DYN and LSI among the small-cap stocks). The differences between M1 versus M2 and M1 versus M3 are, however, slightly higher, but still less than 0.22% for all large-cap stocks and 0.45% for all small-cap stocks. The results based on RMSD give similar conclusion.

To investigate the effects of diurnal adjustment on intraday volatility we consider half-hour intervals from 9:30 through 16:00. Figure 4 plots the average intraday volatility estimates over all trading days,

---

<sup>8</sup>Most studies in the literature evaluate the diurnal factor  $\phi(\cdot)$  at the starting point  $t_i$  of the interval. As our average price duration is around five minutes, it may be more appropriate to take the mid-point. We will report the robustness of our empirical results when the starting value of the interval is used to evaluate the diurnal factor instead.

with the mid-point of each half-hour interval identified in the  $x$ -axis. It can be seen that M2 and M3, which correct for intraday periodicity, are quite close to each other. However, the volatility estimated for the last half-hour is higher for the TT method (M2) than the DA method (M3) for all stocks. Both M2 and M3 have more prominent intraday volatility smile (a deeper U-shape) than the volatility estimate with no periodicity correction (M1). Furthermore, while the lowest intraday volatility for M2 and M3 occurs between 12:45 and 13:15, the lowest intraday volatility for M1 occurs at 13:45 for almost all stocks.<sup>9</sup> As the average duration curves in Figure 2 peak around 13:00, the intraday volatility curve using M1 appears to be biased.

To examine the robustness of the results, we vary the treatment of multiple trades for M2 and M3, as well as the evaluation of the diurnal factor  $\phi(\cdot)$ . First, when the same treatment of multiple trades is applied to the two periodicity adjustment methods, their results are very similar and almost undistinguishable. Thus, the difference in the average volatility at the close of the market is due to the treatment of multiple trades rather than the choice of the method of adjustment. Second, if we evaluate  $\phi(\cdot)$  in M3 at the starting point of the interval rather than the mid-point, M2 is found to have a slightly deeper volatility smile than M3, even when the same treatment of multiple trades is applied. Thus, M3 is not insensitive to the choice of the time point of correction. This is an important drawback of the M3 method and is generally overlooked in the literature.

We further investigate the autocorrelation function (ACF) of 5-minute absolute returns with or without intraday periodicity and volatility adjustment. We follow the methodology of Andersen and Bollerslev (1997) with some modifications. First, we compute the differenced logarithmic price over 5-minute (300-second) intervals, denoted by  $r_i = \log p_{i+1} - \log p_i$ , where  $p_i$  is the price at time  $t_i$ , and calculate the absolute raw return (ARR)  $|r_i/300|$ . Next, we compute the absolute 5-minute return scaled by diurnally transformed duration, i.e.,  $|r_i/(\tilde{t}_{i+1} - \tilde{t}_i)|$ , and call this the absolute filtered return (AFR). Finally, we adjust the return by its corresponding volatility estimate. This adjustment is computed by two methods. Denoting  $\hat{\sigma}$  as the return standard deviation of the day computed using M2, we calculate the absolute standardized return as  $|r_i/[\hat{\sigma}(\tilde{t}_{i+1} - \tilde{t}_i)]|$ , and denote this method by ASR1. In addition, we compute the return standard deviation over the 5-minute interval  $(t_i, t_{i+1})$  using the M2 method and denote it by  $\hat{\sigma}_i$ . We then calculate the absolute standardized return as  $|r_i/\hat{\sigma}_i|$  and denote this method by ASR2.

Figure 5 plots the ACF of the stocks up to 1 day lag ( $12 \times 6.5 = 78$  lags). It can be seen that the ACF of the absolute raw return has a very strong U-shape. The ACF of the absolute filtered return, however, is much flatter, which demonstrates that much of the intraday periodicity in correlation has been removed. While the ACF of the absolute filtered return is still very slowly declining, those of the absolute standardized returns die down to zero much faster, suggesting that the long memory feature of the filtered return is due to the failure in correcting for the clustering of volatility. Figure 6 plots

---

<sup>9</sup>This regularity applies to all 20 stocks in this study (see the accompanying Appendix of Figures).

the ACF of the stocks up to 5 days (390 lags). The periodicity of the absolute raw return is very clear, which is due to the intraday volatility smile as shown in Figure 4. While the periodicity of the absolute filtered return and absolute standardized returns is much reduced, there are still peaks over cycles of about 78 lags. The peaks are particularly prominent for the small-cap stock BMS.

Table 3 summarizes the serial correlation coefficients up to 5 lags as well as the Ljung-Box statistics for all 20 stocks. Due to the large sample size, which varies from 57642 to 58344, the Ljung-Box statistics are highly significant for all residuals. It can be observed, however, that their values are declining in the order of ARR, AFR, ASR1 and ASR2. In particular, the serial correlation coefficients of ASR2 are generally quite small, indicating that the use of the 5-minute volatility estimate has removed much of the periodicity, although it has not been as successful for the small-cap stocks as for the large-cap stocks.

### 4.3 Monte Carlo Simulation of Transaction Data

We now examine if data simulated based on the estimated PACD model can replicate the intraday periodicity. We use the estimated model for the IBM and BMS stocks on TT-adjusted data to simulate price durations and thus the occurrence of the price events in diurnally transformed time. We then transform these times to calendar times and calculate the average duration, which are shown in Figure 7. It can be seen that the intraday average duration graphs mimic quite well the stylized intraday periodicity in duration, although for the BMS stock the drop in average duration at the early part of the trade does not show up in the simulation.<sup>10</sup> In contrast, the simulated path based on the estimated PACD model for raw duration (hence calendar times are directly simulated) shows it is not able to generate the duration periodicity commonly documented. This is despite the fact that the daily volatility estimates using M1 appear to work well.

## 5 Conclusion

In this paper we investigate the use of two methods in dealing with intraday periodicity of high-frequency stock price data, the Duration Adjustment (DA) method and the Time Transformation (TT) method, for the estimation of intraday volatility. The DA method diurnally adjust the raw durations by estimating a diurnal factor  $\phi(\cdot)$ , while the TT method transforms the calendar time to a diurnally adjusted time from which adjusted duration is calculated.

Our empirical results show that whether or not the duration data are adjusted for intraday periodicity has little impact on the daily volatility estimates using the ACD-ICV method. However, the intraday volatility smile shows a more prominent U-shape when the duration data are corrected for intraday periodicity. An important issue in the implementation of the adjustment is how multiple trades at the same time stamp are treated. Intraday volatility smiles are more prominent if multiple trades

---

<sup>10</sup>Note that the average price duration in Figure 7 are for the price events defined by a given price range  $\delta$ . They are different from the plots in Figure 2, which are based on all transactions.

are treated as separate trades rather than one trade, regardless of the method of adjustment. Thus, deleting observations with zero duration dampens the effects of intraday periodicity and may lead to under-estimated intraday volatility in periods of high trading activities.

While the DA method is widely used in the literature, the TT method has several clear advantages. The theoretical motivation of the TT method is easy to understand and its empirical implementation is straightforward. Also, the results of the DA method is not robust to the point of evaluation of the diurnal factor. For studies that require simulating duration data, such as for the estimation of intraday value at risk, the TT method is simple to use.

## References

1. Andersen, T. G., and T. Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance*, 4, 115 – 158.
2. Barndorff-Nielsen, O.E., P. R. Hansen, A. Lunde, and N. Shephard, 2009, Realized kernels in practice: trades and quotes, *Econometrics Journal*, 12, C1-C32.
3. Bauwens, L., and P. Giot, 2000, The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks, *Annals d' Économie et de Statistique*, 60, 117 – 149.
4. Dionne, G., P. Duschene, and M. Pacurar, 2009, Intraday value at risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange, *Journal of Empirical Finance*, 16, 777 – 792.
5. Drost, F. C., and B. J. M. Werker, 2004, Semiparametric duration models, *Journal of Business and Economic Statistics*, 22, 40 – 50.
6. Dufour, A., and R. F. Engle, 2000, Time and the impact of a trade, *Journal of Finance*, 55, 2467 – 2498.
7. Engle, R. F., and J. R. Russell, 1998, Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica*, 66, 1127 – 1162.
8. Fernandes, M., and J. Grammig, 2006, A family of autoregressive conditional duration models, *Journal of Econometrics*, 130, 1 – 23.
9. Giot, P., 2005, Market risk models for intraday data, *European Journal of Finance*, 11, 309 – 324.
10. Rodríguez-Poo, J. M., D. Veredas, and A. Espasa, 2007, Semiparametric estimation for financial durations, in *High Frequency Financial Econometrics: Recent Developments*, Eds L. Bauwens, W. Polhmeier, and D. Veredas, Springer.
11. Tsay, R. S., 2002, *Analysis of Financial Times Series*, Wiley.

12. Tse, Y. K., and T. T. Yang, 2012, Estimation of high-frequency volatility: An autoregressive conditional duration approach, working paper, Singapore Management University, forthcoming in Journal of Business and Economic Statistics.
13. Wu, Z., 2011, On the intraday periodicity duration adjustment of high-frequency data, working paper, National University of Singapore, forthcoming in Journal of Empirical Finance.

## Appendix

### A.1 Implementation of the Duration Adjustment Method

We divide each trading day (9:30 – 16:00) into 14 intervals, with the first and last interval being 15 minutes each and all others being 30 minutes each. Let the number of trading days in the data set be  $T$  and the number of trades in the  $k$ th time interval on day  $i$  be  $N_{ik}$ , where  $i = 1, \dots, T$  and  $k = 1, \dots, 14$ . For the purpose of computing  $N_{ik}$ , we may take multiple trades at the same time stamp as one trade or separate trades. Denote  $J_k$  as the number of seconds in the  $k$ th time interval, e.g.,  $J_1 = 900$  and  $J_2 = 1800$ . When multiple trades at the same time stamp are taken as one trade, we have

$$N_{ik} = \sum_{j=1}^{J_k} \mathbf{1}(N_{ikj} > 0),$$

where  $\mathbf{1}(N_{ikj} > 0)$  is the indicator function taking value 1 when  $N_{ikj} > 0$  and 0 otherwise, with  $N_{ikj}$  being the number of trades at the  $j$ th second in the  $k$ th interval on the  $i$ th trading day. However, when multiple trades at the same time stamp are taken as separate trades, we have

$$N_{ik} = \sum_{j=1}^{J_k} N_{ikj}.$$

For the  $k$ th time interval, the average number of trades over all trading days is computed as

$$N_k = \frac{1}{T} \sum_{i=1}^T N_{ik}.$$

We then calculate the average trade duration in the  $k$ th time interval as  $\bar{x}_k = J_k/N_k$ . Let  $t_k$  be the mid-point of the  $k$ th time interval. We set the diurnal factor at the mid-point of each time interval to be equal to the standardized average trade duration of the interval so that

$$\phi(t_k) = \frac{\bar{x}_k}{\sum_{k=1}^{14} \bar{x}_k}.$$

Finally, the diurnal factor function  $\phi(t)$  is calculated using the cubic spline smoothing with  $\phi(t_k)$  as the knots, for  $k = 1, \dots, 14$ .

### A.2 Implementation of the Time Transformation Method

Let the strictly increasing time points  $t_0, t_1, \dots, t_{23400}$  separately denote the time 9:30:00, 9:30:01,  $\dots$ , 16:00:00 in second, so that  $t_0 = 0$ ,  $t_1 = 1$ ,  $\dots$ ,  $t_{23400} = 23400$ . For any two calendar-time points  $t_i$  and

$t_j$  ( $t_i < t_j$ ) the corresponding calendar-time duration is  $x_{ij} = t_j - t_i$ . Let  $n_i$  denote the number of trades at time  $t_i$  aggregated over all days, for  $i = 1, \dots, 23400$ . Again for the purpose of computing  $n_i$ , we may take multiple trades at the same time stamp as one trade or separate trades. We compute

$$N_{t_k} = \sum_{i=1}^k n_i, \quad k = 1, \dots, 23400,$$

and

$$N_T = \sum_{i=1}^{23400} n_i.$$

Thus,  $N_{t_k}$  is the total number of trades up to and including time  $t_k$  aggregated over all trading days in the sample. This function may be further smoothed by linear interpolation at the neighborhood of  $t_k$  for some  $k$  when  $n_k = 0$ . The time-transformation function  $Q(t_k)$  is then computed as

$$Q(t_k) = \frac{N_{t_k}}{N_T}, \quad k = 0, 1, \dots, 23400,$$

with  $N_{t_0} = 0$ , so that

$$\tilde{t}_k = 23400 Q(t_k) = 23400 \left[ \frac{N_{t_k}}{N_T} \right], \quad k = 0, 1, \dots, 23400,$$

where  $\tilde{t}_k$  is the diurnally transformed time corresponding to calendar time  $t_k$ .

**Table 1:** Summary statistics of stocks

<u>Panel A: Large-cap stocks</u>										
	CVX	GE	IBM	JNJ	JPM	PFE	PG	T	WMT	XOM
Number of days	748	748	748	747	747	747	748	740	748	748
<u>Multiple trades at the same time treated as separate trades</u>										
Number of transactions (million)	7.709	7.107	5.919	5.498	6.544	6.305	5.598	5.647	6.280	10.914
Avg transactions per day	10306	9501	7913	7360	8760	8440	7484	7631	8396	14591
Avg duration (second)	2.27	2.46	2.96	3.18	2.67	2.77	3.13	3.07	2.79	1.60
<u>Multiple trades at the same time treated as one trade</u>										
Number of transactions (million)	4.898	4.492	4.006	3.807	4.109	4.352	3.721	3.512	4.289	6.181
Avg transactions per day	6548	6005	5356	5096	5501	5826	4975	4746	5735	8264
Avg duration (second)	3.57	3.90	4.37	4.59	4.25	4.02	4.70	4.93	4.08	2.83
<u>Panel B: Small-cap stocks</u>										
	BMS	CMS	DYN	FII	KG	LEG	LSI	MYL	PLL	TE
Number of days	747	743	744	745	743	746	743	747	745	739
<u>Multiple trades at the same time treated as separate trades</u>										
Number of transactions (million)	1.080	1.529	1.509	1.273	1.694	1.385	2.266	1.717	1.264	1.161
Avg transactions per day	1446	2058	2028	1708	2280	1856	3050	2298	1697	1571
Avg duration (second)	16.20	11.38	11.54	13.71	10.27	12.61	7.68	10.19	13.80	14.90
<u>Multiple trades at the same time treated as one trade</u>										
Number of transactions (million)	0.922	1.307	1.251	1.062	1.411	1.156	1.844	1.404	1.071	0.998
Avg transactions per day	1235	1759	1681	1425	1899	1549	2482	1880	1438	1351
Avg duration (second)	18.97	13.31	13.93	16.43	12.33	15.11	9.43	12.45	16.29	17.34

**Table 2:** Results of daily volatility estimation

Stock	Mean Daily ICV			MAD			RMSD		
	M1	M2	M3	M1-M2	M1-M3	M2-M3	M1-M2	M1-M3	M2-M3
Panel A: Large-cap stocks									
CVX	19.1913	19.2013	19.2085	0.1845	0.0969	0.1013	0.2513	0.1296	0.1389
GE	12.3780	12.3861	12.3822	0.1159	0.0720	0.0483	0.1722	0.1126	0.0691
IBM	14.7162	14.7506	14.7408	0.1058	0.0787	0.0376	0.1382	0.1069	0.0497
JNJ	11.1155	11.1476	11.1371	0.0839	0.0618	0.0286	0.1080	0.0800	0.0377
JPM	16.3696	16.3481	16.3473	0.2148	0.1562	0.0635	0.2988	0.2224	0.0863
PFE	16.1752	16.2178	16.2047	0.0962	0.0778	0.0295	0.1181	0.0966	0.0373
PG	12.8189	12.8722	12.8547	0.1021	0.0799	0.0313	0.1275	0.1014	0.0405
T	16.0181	16.0427	16.0366	0.1518	0.0963	0.0712	0.2302	0.1522	0.1025
WMT	14.9632	15.0001	14.9904	0.0894	0.0629	0.0389	0.1146	0.0813	0.0498
XOM	17.9755	17.9809	17.9752	0.1744	0.0881	0.0923	0.2236	0.1158	0.1170
Panel B: Small-cap stocks									
BMS	15.3549	15.3787	15.3940	0.1110	0.1191	0.0612	0.1466	0.1584	0.0806
CMS	19.1475	19.2210	19.2027	0.2026	0.1623	0.0796	0.2597	0.2392	0.1318
DYN	30.1204	30.1279	30.1192	0.3705	0.3206	0.1007	0.5074	0.4358	0.1368
FII	15.7800	15.8294	15.8343	0.1360	0.1454	0.0530	0.1991	0.2004	0.0800
KG	20.6878	20.7462	20.7379	0.1802	0.1507	0.0703	0.2389	0.2031	0.0987
LEG	17.5914	17.6685	17.6491	0.2006	0.1422	0.0740	0.2537	0.1821	0.0974
LSI	29.8013	29.9719	29.9387	0.4498	0.3660	0.1045	0.5596	0.4559	0.1384
MYL	18.7517	18.7783	18.7809	0.1617	0.1311	0.0793	0.2212	0.1822	0.1165
PLL	15.5531	15.5928	15.5952	0.1086	0.0977	0.0472	0.1545	0.1341	0.0685
TE	15.9287	15.9832	15.9723	0.1616	0.1469	0.0518	0.2214	0.2187	0.0825

**Notes:** Volatility is annualized standard deviation in percentage. MAD is mean absolute deviation. RMSD is Root mean-squared deviation. M1 uses raw duration without diurnal adjustment. M2 uses diurnalized duration based on the TT method. M3 uses diurnalized duration based on the DA method.



**Table 3:** ACF and Ljung-Box statistics for raw and adjusted returns

Panel A: Large-cap stocks											
		CVX	GE	IBM	JNJ	JPM	PFE	PG	T	WMT	XOM
ARR	LB-Stat	21302	23986	23197	17207	50958	14123	17424	21988	19304	18731
	$\hat{\rho}_1$	0.2224	0.2372	0.2696	0.2636	0.3084	0.2233	0.2404	0.2636	0.2486	0.1973
	$\hat{\rho}_2$	0.1946	0.2123	0.2202	0.1972	0.2853	0.1833	0.1901	0.2158	0.2068	0.1783
	$\hat{\rho}_3$	0.1770	0.1934	0.1947	0.1730	0.2707	0.1733	0.1851	0.1860	0.1883	0.1681
	$\hat{\rho}_4$	0.1695	0.1819	0.1875	0.1660	0.2690	0.1536	0.1686	0.1848	0.1772	0.1667
	$\hat{\rho}_5$	0.1616	0.1770	0.1777	0.1495	0.2502	0.1463	0.1528	0.1757	0.1695	0.1449
AFR	LB-Stat	20029	18267	18016	12293	48461	10053	11950	18420	14307	17721
	$\hat{\rho}_1$	0.2030	0.2021	0.2318	0.2236	0.2828	0.1951	0.1949	0.2281	0.2086	0.1841
	$\hat{\rho}_2$	0.1834	0.1804	0.1821	0.1570	0.2641	0.1508	0.1462	0.1861	0.1685	0.1684
	$\hat{\rho}_3$	0.1653	0.1570	0.1607	0.1393	0.2490	0.1408	0.1458	0.1620	0.1505	0.1570
	$\hat{\rho}_4$	0.1575	0.1498	0.1556	0.1329	0.2560	0.1229	0.1293	0.1611	0.1468	0.1556
	$\hat{\rho}_5$	0.1517	0.1453	0.1484	0.1190	0.2393	0.1189	0.1200	0.1557	0.1384	0.1389
ASR1	LB-Stat	2292	2325	2593	3843	2583	3376	3584	3153	3146	1623
	$\hat{\rho}_1$	0.1101	0.1166	0.1362	0.1691	0.1180	0.1453	0.1449	0.1464	0.1390	0.0850
	$\hat{\rho}_2$	0.0785	0.0840	0.0878	0.1006	0.0829	0.1001	0.0929	0.0942	0.0942	0.0670
	$\hat{\rho}_3$	0.0648	0.0620	0.0667	0.0801	0.0704	0.0889	0.0889	0.0700	0.0786	0.0612
	$\hat{\rho}_4$	0.0559	0.0535	0.0590	0.0720	0.0648	0.0720	0.0708	0.0697	0.0716	0.0531
	$\hat{\rho}_5$	0.0487	0.0501	0.0492	0.0594	0.0560	0.0632	0.0603	0.0607	0.0628	0.0385
ASR2	LB-Stat	754	950	940	1436	983	1402	1292	1312	1056	423
	$\hat{\rho}_1$	0.0831	0.0961	0.0956	0.1229	0.0929	0.1057	0.1052	0.1141	0.1002	0.0626
	$\hat{\rho}_2$	0.0449	0.0568	0.0486	0.0550	0.0522	0.0588	0.0498	0.0561	0.0533	0.0353
	$\hat{\rho}_3$	0.0311	0.0351	0.0328	0.0373	0.0411	0.0525	0.0493	0.0344	0.0369	0.0272
	$\hat{\rho}_4$	0.0243	0.0247	0.0281	0.0354	0.0317	0.0395	0.0361	0.0371	0.0318	0.0216
	$\hat{\rho}_5$	0.0188	0.0219	0.0240	0.0254	0.0259	0.0352	0.0278	0.0317	0.0265	0.0091
Panel B: Small-cap stocks											
		BMS	CMS	DYN	FII	KG	LEG	LSI	MYL	PLL	TE
ARR	LB-Stat	20383	18043	10121	29855	24105	16654	14831	23827	23711	20079
	$\hat{\rho}_1$	0.2967	0.2719	0.2282	0.2971	0.3147	0.2780	0.2617	0.3067	0.3060	0.2701
	$\hat{\rho}_2$	0.2329	0.2096	0.1672	0.2476	0.2470	0.2051	0.2105	0.2401	0.2550	0.2181
	$\hat{\rho}_3$	0.1930	0.1784	0.1378	0.2284	0.2129	0.1822	0.1792	0.2227	0.2330	0.1909
	$\hat{\rho}_4$	0.1770	0.1647	0.1181	0.2059	0.1878	0.1649	0.1420	0.1908	0.1904	0.1697
	$\hat{\rho}_5$	0.1578	0.1533	0.1073	0.1987	0.1654	0.1514	0.1315	0.1788	0.1785	0.1623
AFR	LB-Stat	7193	10660	6518	15505	16431	7368	9000	19657	13816	13358
	$\hat{\rho}_1$	0.1949	0.2099	0.1905	0.2507	0.2602	0.2172	0.1973	0.2745	0.2411	0.2197
	$\hat{\rho}_2$	0.1421	0.1481	0.1292	0.1799	0.1965	0.1304	0.1545	0.2081	0.1832	0.1662
	$\hat{\rho}_3$	0.1081	0.1270	0.1033	0.1601	0.1653	0.1100	0.1292	0.1912	0.1705	0.1499
	$\hat{\rho}_4$	0.0957	0.1165	0.0886	0.1400	0.1449	0.0965	0.0972	0.1704	0.1414	0.1301
	$\hat{\rho}_5$	0.0858	0.1150	0.0795	0.1309	0.1291	0.0866	0.0926	0.1607	0.1316	0.1237
ASR1	LB-Stat	2308	3283	3979	2900	4238	3022	4223	4820	3148	3580
	$\hat{\rho}_1$	0.1519	0.1570	0.1581	0.1726	0.1840	0.1795	0.1614	0.1874	0.1697	0.1634
	$\hat{\rho}_2$	0.0876	0.0928	0.0968	0.0964	0.1174	0.0906	0.1181	0.1181	0.1054	0.1059
	$\hat{\rho}_3$	0.0552	0.0778	0.0808	0.0698	0.0900	0.0676	0.0918	0.0998	0.0829	0.0900
	$\hat{\rho}_4$	0.0421	0.0606	0.0698	0.0480	0.0726	0.0546	0.0605	0.0819	0.0585	0.0628
	$\hat{\rho}_5$	0.0370	0.0595	0.0621	0.0407	0.0524	0.0397	0.0597	0.0670	0.0467	0.0525
ASR2	LB-Stat	1365	1904	3093	1227	1942	1131	2210	2021	1259	1931
	$\hat{\rho}_1$	0.1192	0.1200	0.1476	0.1182	0.1324	0.1121	0.1280	0.1327	0.1188	0.1247
	$\hat{\rho}_2$	0.0589	0.0667	0.0846	0.0589	0.0765	0.0529	0.0801	0.0723	0.0604	0.0728
	$\hat{\rho}_3$	0.0344	0.0553	0.0704	0.0369	0.0553	0.0332	0.0591	0.0609	0.0457	0.0615
	$\hat{\rho}_4$	0.0306	0.0403	0.0582	0.0215	0.0448	0.0304	0.0328	0.0493	0.0297	0.0407
	$\hat{\rho}_5$	0.0280	0.0403	0.0499	0.0239	0.0308	0.0201	0.0380	0.0377	0.0241	0.0368

**Notes:** ARR is the absolute raw return, AFR is the absolute filtered return, ASR1 is the absolute return standardized by daily standard deviation scaled by transformed duration, and ASR2 is the absolute return standardized by 5-minute standard deviation. LB-Stat is the Ljung-Box statistic with 15 lags.

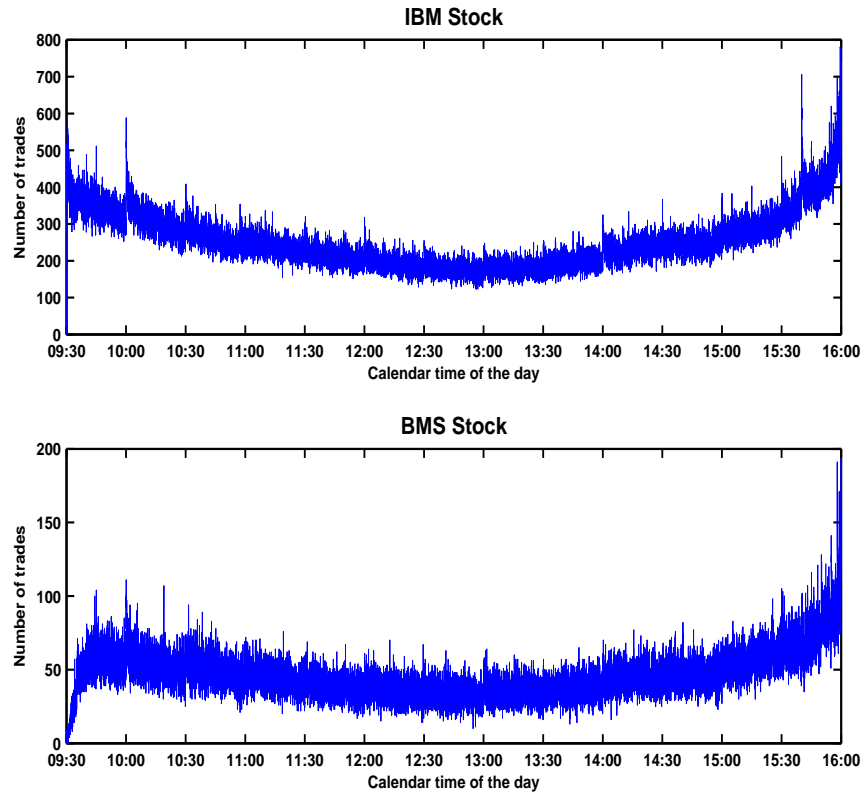


Figure 1: Number of trades at each second over all trading days in sample period.

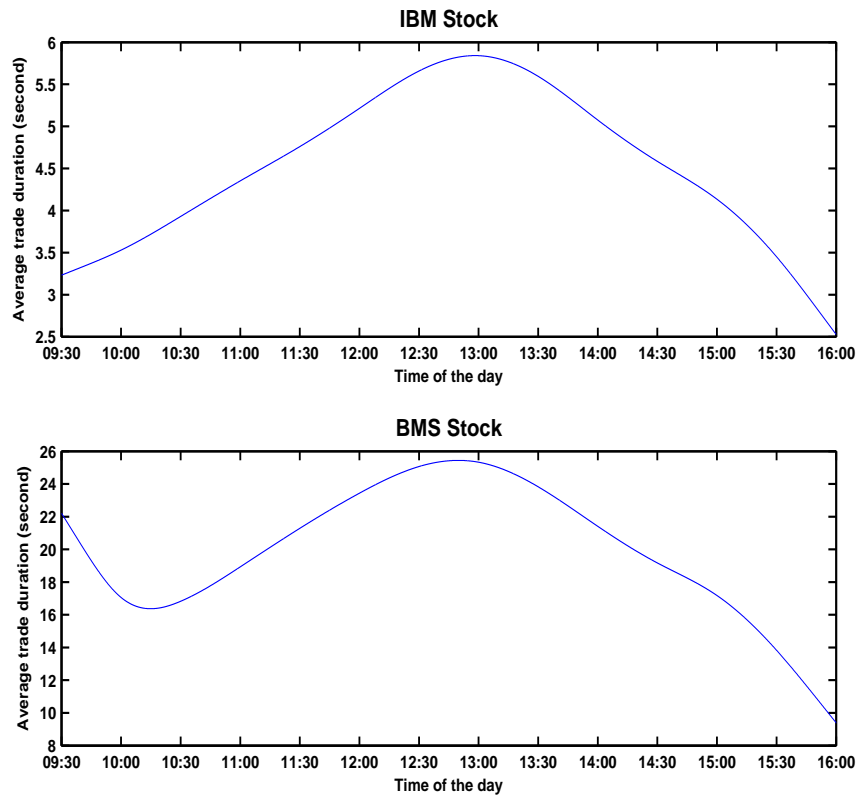


Figure 2: Smoothed average duration by time of the day over all trading days in sample period.

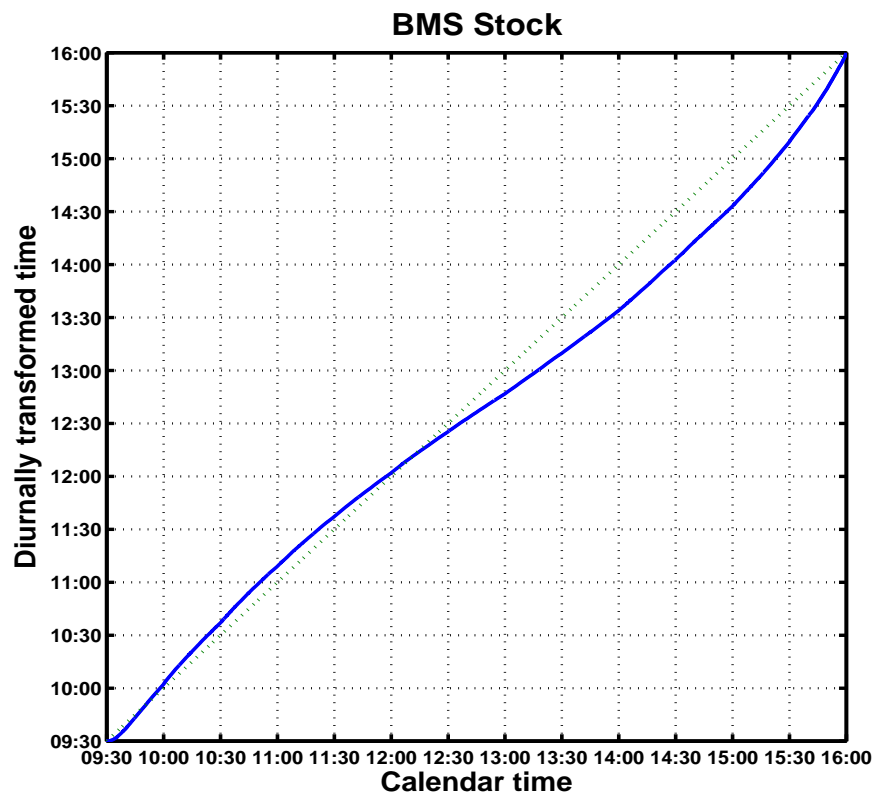
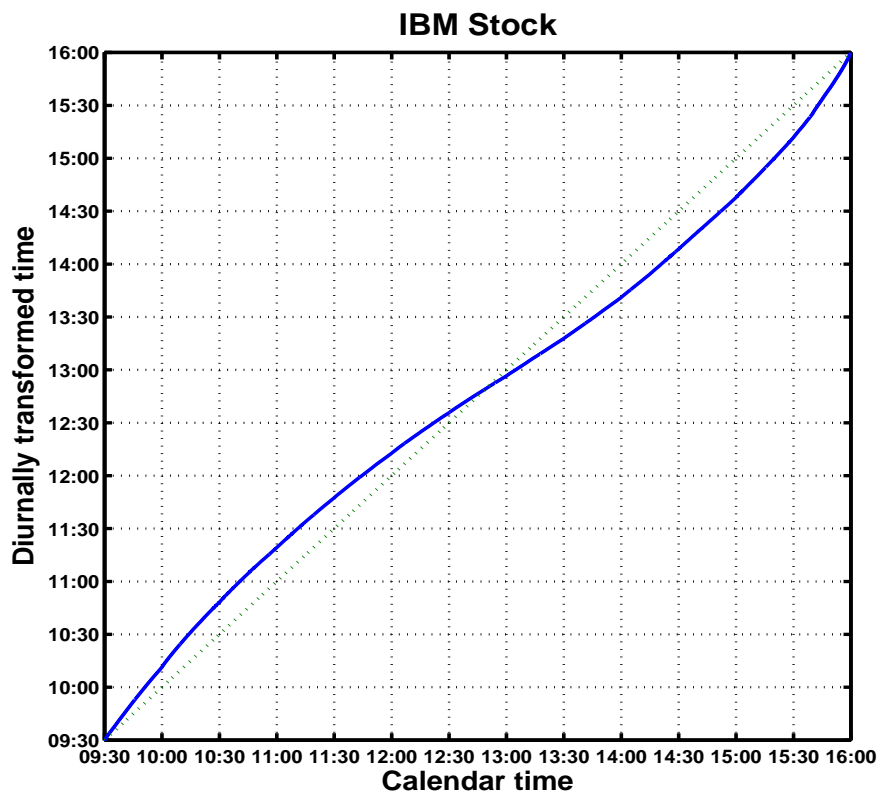


Figure 3: Diurnally transformed time versus calendar time.

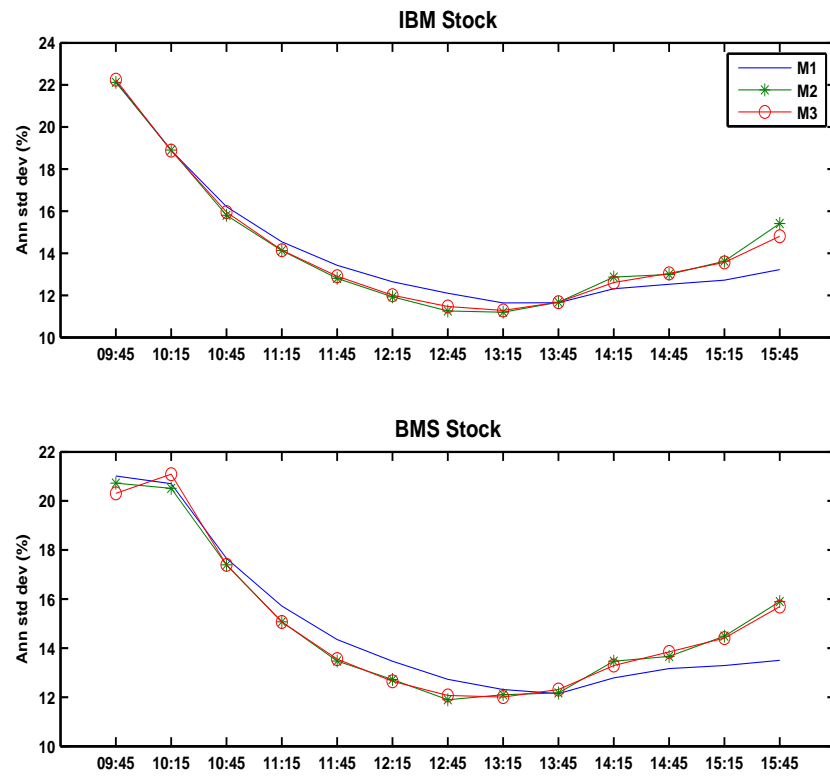


Figure 4: Intraday volatility calculated using M1, M2 and M3.

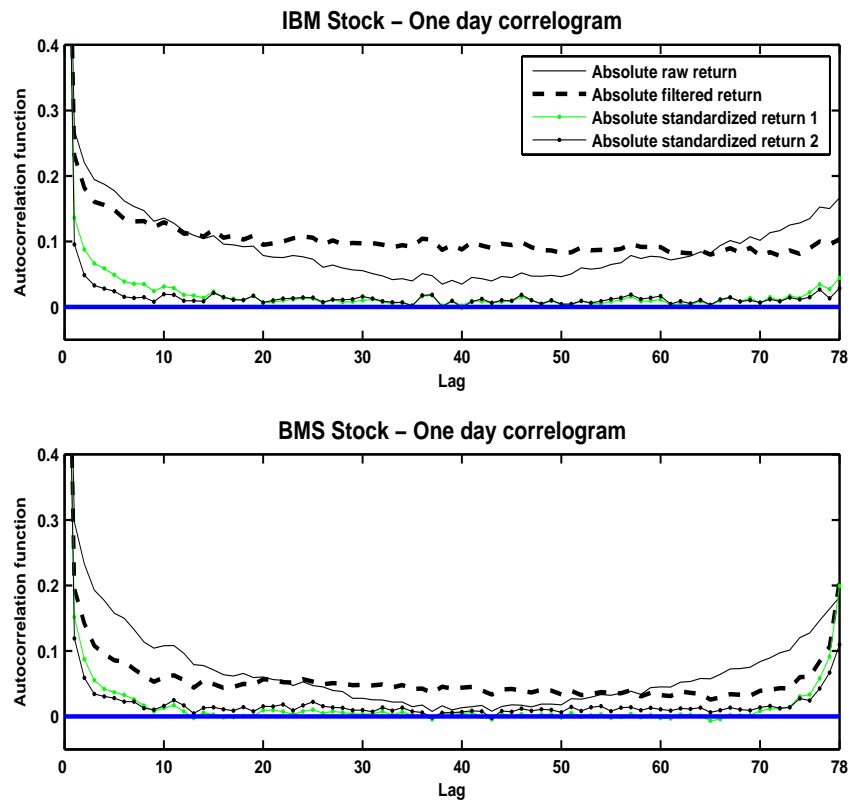


Figure 5: ACF of 5-minute absolute returns over one day.

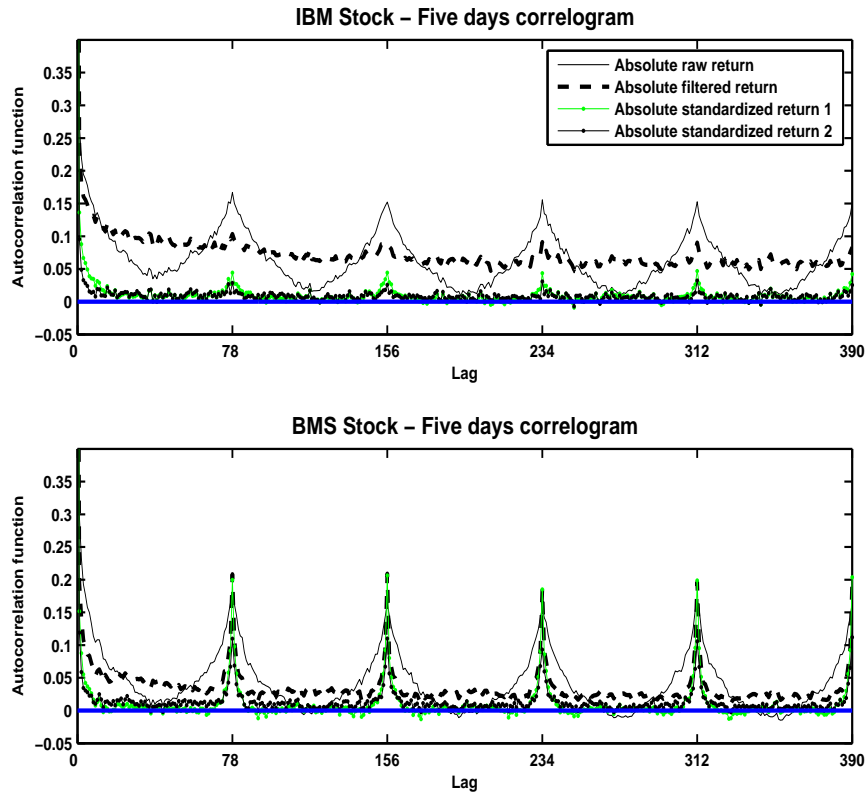


Figure 6: ACF of 5-minute absolute returns over five days.

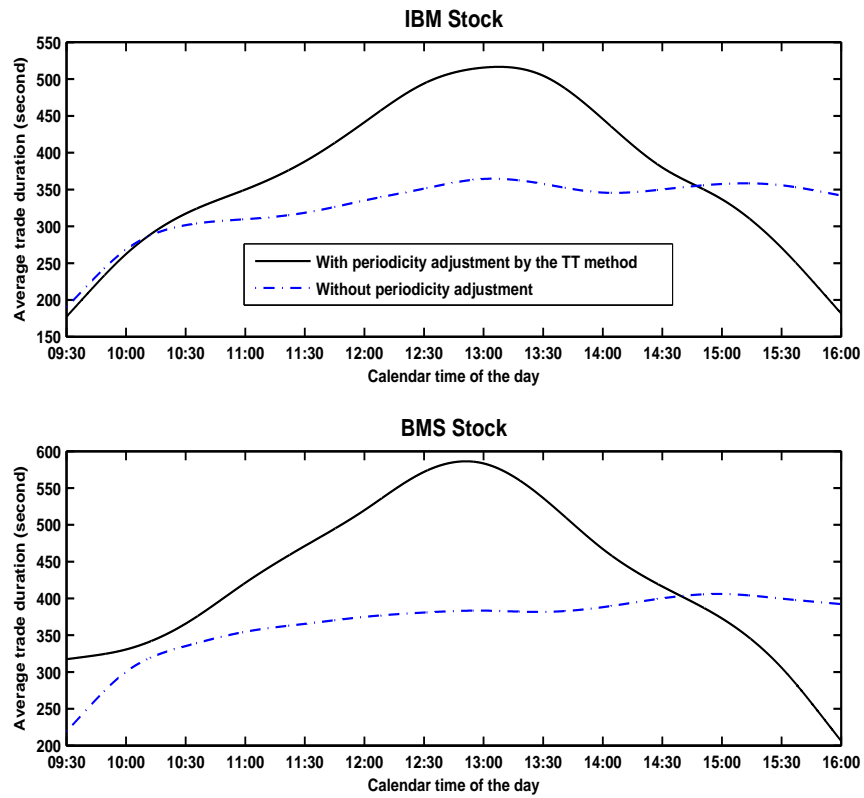


Figure 7: Average simulated price duration.